Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators via characterizations of operator concave functions (Recent Topics on Operator inequalities)

Furuta, Takayuki

数理解析研究所講究録 1359 巻 2004 年 85-92

Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators via characterizations of operator concave functions

東京理科大学理学部　古田孝之 (Takayuki Furuta)

Abstract. We shall state the following parametric extensions of Shannon inequality and its reverse one in Hilbert space operators. *Let $p \in [0,1]$ and also let $\{A_1, A_2, ..., A_n\}$ and $\{B_1, B_2, ..., B_n\}$ be two sequences of strictly positive operators on a Hilbert space H such that $\sum_{j=1}^{n} A_j \sharp_p B_j \leq I$. Then*

$$\sum_{j=1}^{n} S_{p+1}(A_j|B_j) \geq \Big[\sum_{j=1}^{n}(A_j \natural_{p+1} B_j) + (I - \sum_{j=1}^{n} A_j \sharp_p B_j)\Big] \log\Big[\sum_{j=1}^{n}(A_j \natural_{p+1} B_j) + (I - \sum_{j=1}^{n} A_j \sharp_p B_j)\Big]$$
$$\geq \log\Big[\sum_{j=1}^{n}(A_j \natural_{p+1} B_j) + (I - \sum_{j=1}^{n} A_j \sharp_p B_j)\Big] \geq \sum_{j=1}^{n} S_p(A_j|B_j) \geq -\log\Big[\sum_{j=1}^{n}(A_j \natural_{p-1} B_j) + (I - \sum_{j=1}^{n} A_j \sharp_p B_j)\Big]$$
$$\geq -\Big[\sum_{j=1}^{n}(A_j \natural_{p-1} B_j) + (I - \sum_{j=1}^{n} A_j \sharp_p B_j)\Big] \log\Big[\sum_{j=1}^{n}(A_j \natural_{p-1} B_j) + (I - \sum_{j=1}^{n} A_j \sharp_p B_j)\Big] \geq \sum_{j=1}^{n} S_{p-1}(A_j|B_j)$$

where $S_q(A|B) = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^q(\log A^{\frac{-1}{2}}BA^{\frac{-1}{2}})A^{\frac{1}{2}}$ for $A > 0$, $B > 0$ and any real number q and $A \natural_q B = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^q A^{\frac{1}{2}}$ for $A > 0$, $B > 0$ and any real number q.

In particular, if $\sum_{j=1}^{n} A_j = \sum_{j=1}^{n} B_j = I$, then

$$\sum_{j=1}^{n} S_2(A_j|B_j) \geq \Big[\sum_{j=1}^{n} B_j A_j^{-1} B_j\Big] \log\Big[\sum_{j=1}^{n} B_j A_j^{-1} B_j\Big] \geq \log\Big[\sum_{j=1}^{n} B_j A_j^{-1} B_j\Big] \geq \sum_{j=1}^{n} S_1(A_j|B_j) \geq 0$$
$$\geq \sum_{j=1}^{n} S(A_j|B_j) \geq -\log\Big[\sum_{j=1}^{n} A_j B_j^{-1} A_j\Big] \geq -\Big[\sum_{j=1}^{n} A_j B_j^{-1} A_j\Big] \log\Big[\sum_{j=1}^{n} A_j B_j^{-1} A_j\Big] \geq \sum_{j=1}^{n} S_{-1}(A_j|B_j)$$

where $S(A|B) = S_0(A|B) = A^{\frac{1}{2}}(\log A^{\frac{-1}{2}}BA^{\frac{-1}{2}})A^{\frac{1}{2}}$ which is the relative operator entropy of $A > 0$ and $B > 0$.

Our results can be considered as parametric extensions of the following celebrated Shannon inequality ([7],[9] and [233 p ,1]) which is very useful and so famous in information theory. *Let $\{a_1, a_2, ..., a_n\}$ and $\{b_1, b_2, ..., b_n\}$ be two probability vectors. Then*

$$0 \geq \sum_{j=1}^{n} a_j \log b_j - \sum_{j=1}^{n} a_j \log a_j$$ (see inequalities (2.4) of Corollary 2.4).

§1 Introduction

First the Shannon inequality asserts: *Let* $\{a_1, a_2, ..., a_n\}$ *and* $\{b_1, b_2, ..., b_n\}$ *be two probability vectors. Then*

$$0 \geq \sum_{j=1}^{n} a_j \log \frac{b_j}{a_j}. \tag{1.1}$$

We remark that $0 \geq \sum_{j=1}^{n} a_j \log \frac{b_j}{a_j}$ in (1.1) is equivalent to $D = \sum_{j=1}^{n} a_j \log \frac{a_j}{b_j} \geq 0$ which is the original number type Shannon inequality and this D is called "divergence" in [7] and [9].

In this paper we shall state parametric extensions of Shannon inequality and its reverse one in Hilbert space operators.

A bounded linear operator T on a Hilbert space H is said to be positive (denoted by $T \geq 0$) if $(Tx, x) \geq 0$ for all $x \in H$ and also an operator T is said to be strictly positive (denoted by $T > 0$) if T is invertible and positive.

Definition 1.1. $S_q(A|B)$ for $A > 0$, $B > 0$ and any real number q is defined by

$$S_q(A|B) = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^q(\log A^{\frac{-1}{2}}BA^{\frac{-1}{2}})A^{\frac{1}{2}}.$$

We recall that $S_0(A|B) = A^{\frac{1}{2}}(\log A^{\frac{-1}{2}}BA^{\frac{-1}{2}})A^{\frac{1}{2}} = S(A|B)$ is the relative operator entropy in [2] and $S(A|I) = -A\log A$ is the usual operator entropy in [8].

Definition 1.2. *$A\natural_q B$ for $A > 0$ and $B >$ and any real number q is defined by*

$$A\natural_q B = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^q A^{\frac{1}{2}}$$

and $A\natural_p B$ for $p \in [0,1]$ just coincides with $A\sharp_p B$ which is well known as p-power mean.

We remark that $S_1(A|B) = -S(B|A)$ and moreover $S_q(A|B) = -S_{1-q}(B|A)$ for any q. Following after Definition 1.1, The original Shannon inequality can be expressed as follows:

$$0 \geq \sum_{j=1}^{n} a_j \log \frac{b_j}{a_j} = \sum_{j=1}^{n} a_j^{\frac{1}{2}}(\log a_j^{\frac{-1}{2}} b_j a_j^{\frac{-1}{2}})a_j^{\frac{1}{2}} = \sum_{j=1}^{n} S(a_j|b_j).$$

Consequently $0 \geq \sum_{j=1}^{n} S(a_j|b_j)$ in the original Shannon inequality can be extented to $0 \geq \sum_{j=1}^{n} S(A_j|B_j)$ in operator version case (2.4) of Corollary 2.4, so that the form of (1.1) is convenient for operator type extension. We can summarize the following contrast:

The original Shannon inequality and its reverse one

$$0 \ge \sum_{j=1}^{n} a_j \log \frac{b_j}{a_j} \ge -\log \sum_{j=1}^{n} \frac{a_j^2}{b_j}.$$

for $a_j, b_j > 0$ *with* $1 = \sum_{j=1}^{n} a_j = \sum_{j=1}^{n} b_j$.

The operator version Shannon inequality and its reverse one

$$0 \ge \sum_{j=1}^{n} S(A_j|B_j) \ge -\log \sum_{j=1}^{n} A_j B_j^{-1} A_j.$$

for $A_j, B_j > 0$ with $I = \sum_{j=1}^{n} A_j = \sum_{j=1}^{n} B_j$.

§2 Parametric extensions of operator reverse type Shannon inequality derived from two operator concave functions $f_1(t) = \log t$ and $f_2(t) = -t\log t$

Firstly we shall state the following parametric extensions of Shannon inequality and its reverse one in Hilbert space operators derived from an operator concave function $f(t) = \log t$.

Theorem 2.1. *Let $p \in [0,1]$ and also let $\{A_1, A_2, ..., A_n\}$ and $\{B_1, B_2, ..., B_n\}$ be two sequences of strictly positive operators on a Hilbert space H such that $\sum_{j=1}^{n} A_j \sharp_p B_j \le I$, where I means the identity operator on H. Then*

$$\begin{aligned} \text{(2.1)}\qquad & \log\Big[\sum_{j=1}^{n}(A_j \natural_{p+1} B_j) + t_0(I - \sum_{j=1}^{n} A_j \sharp_p B_j)\Big] - \log t_0(I - \sum_{j=1}^{n} A_j \sharp_p B_j) \\ & \ge \sum_{j=1}^{n} S_p(A_j|B_j) \\ & \ge -\log\Big[\sum_{j=1}^{n}(A_j \natural_{p-1} B_j) + t_0(I - \sum_{j=1}^{n} A_j \sharp_p B_j)\Big] + \log t_0(I - \sum_{j=1}^{n} A_j \sharp_p B_j) \end{aligned}$$

for fixed real number $t_0 > 0$, where $S_p(A|B)$ is defined in Definition 1.1 and $A \natural_q B$ is defined in Definition 1.2.

Secondly we shall state the following parametric extensions of Shannon inequality and its reverse one in Hilbert space operators derived from an operator concave function $f(t) = -t\log t$.

Theorem 2.2. *Let $p \in [0,1]$ and also let $\{A_1, A_2, ..., A_n\}$ and $\{B_1, B_2, ..., B_n\}$ be two sequences of strictly positive operators on a Hilbert space H such that $\sum_{j=1}^{n} A_j \sharp_p B_j \le I$, where I means the identity operator on H. Then*

$$\text{(2.2)}\qquad \sum_{j=1}^{n} S_{p+1}(A_j|B_j)$$

$$\geq \Big[\sum_{j=1}^{n}(A_j\natural_{p+1}B_j) + t_0(I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]\log\Big[\sum_{j=1}^{n}(A_j\natural_{p+1}B_j) + t_0(I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]$$
$$-t_0 \log t_0(I - \sum_{j=1}^{n} A_j\sharp_p B_j) \quad \textit{for fixed real number } t_0 > 0,$$

and

(2.2') $$\sum_{j=1}^{n} S_{p-1}(A_j|B_j)$$
$$\leq -\Big[\sum_{j=1}^{n}(A_j\natural_{p-1}B_j) + t_0(I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]\log\Big[\sum_{j=1}^{n}(A_j\natural_{p-1}B_j) + t_0(I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]$$
$$+t_0 \log t_0(I - \sum_{j=1}^{n} A_j\sharp_p B_j) \quad \textit{for fixed real number } t_0 > 0,$$

where $S_q(A|B)$ *is defined in Definition 1.1 and* $A\natural_q B$ *is defined in Definition 1.2.*

We shall state the following result which can be shown by combining Theorem 2.1 with Theorem 2.2.

Corollary 2.3. *Let* $p \in [0,1]$ *and also let* $\{A_1, A_2, ..., A_n\}$ *and* $\{B_1, B_2, ..., B_n\}$ *be two sequences of strictly positive operators on a Hilbert space* H *such that* $\sum_{j=1}^{n} A_j\sharp_p B_j \leq I$, *where* I *means the identity operator on* H. *Then*

(2.3) $$\sum_{j=1}^{n} S_{p+1}(A_j|B_j)$$
$$\geq \Big[\sum_{j=1}^{n}(A_j\natural_{p+1}B_j) + (I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]\log\Big[\sum_{j=1}^{n}(A_j\natural_{p+1}B_j) + (I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]$$
$$\geq \log\Big[\sum_{j=1}^{n}(A_j\natural_{p+1}B_j) + (I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]$$
$$\geq \sum_{j=1}^{n} S_p(A_j|B_j)$$
$$\geq -\log\Big[\sum_{j=1}^{n}(A_j\natural_{p-1}B_j) + (I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]$$
$$\geq -\Big[\sum_{j=1}^{n}(A_j\natural_{p-1}B_j) + (I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]\log\Big[\sum_{j=1}^{n}(A_j\natural_{p-1}B_j) + (I - \sum_{j=1}^{n} A_j\sharp_p B_j)\Big]$$
$$\geq \sum_{j=1}^{n} S_{p-1}(A_j|B_j)$$

where $S_q(A|B)$ *is defined in Definition 1.1 and* $A\natural_q B$ *is defined in Definition 1.2.*

Corollary 2.3 easily implies the following result which can be considered as *operator version of Shannon inequality and its reverse one.*

Corollary 2.4. *Let $\{A_1, A_2, ..., A_n\}$ and $\{B_1, B_2, ..., B_n\}$ be two sequences of strictly positive operators on a Hilbert space H. If $\sum_{j=1}^{n} A_j = \sum_{j=1}^{n} B_j = I$, then*

$$
\begin{aligned}
(2.4)\quad \sum_{j=1}^{n} S_2(A_j|B_j) &\geq \Big[\sum_{j=1}^{n} B_j A_j^{-1} B_j\Big] \log\Big[\sum_{j=1}^{n} B_j A_j^{-1} B_j\Big] \geq \log\Big[\sum_{j=1}^{n} B_j A_j^{-1} B_j\Big] \\
&\geq \sum_{j=1}^{n} S_1(A_j|B_j) \geq 0 \geq \sum_{j=1}^{n} S(A_j|B_j) \\
&\geq -\log\Big[\sum_{j=1}^{n} A_j B_j^{-1} A_j\Big] \geq -\Big[\sum_{j=1}^{n} A_j B_j^{-1} A_j\Big] \log\Big[\sum_{j=1}^{n} A_j B_j^{-1} A_j\Big] \\
&\geq \sum_{j=1}^{n} S_{-1}(A_j|B_j).
\end{aligned}
$$

Remark 2.1. We recall $S_q(A|B)$ for $A > 0$, $B > 0$ and any real number q as follows:

$$S_q(A|B) = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^q(\log A^{\frac{-1}{2}}BA^{\frac{-1}{2}})A^{\frac{1}{2}}.$$

By an easy calculation we have

$$\tfrac{d}{dq}[S_q(A|B)] = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^q[\log A^{\frac{-1}{2}}BA^{\frac{-1}{2}}]^2 A^{\frac{1}{2}} \geq 0,$$

so that $S_q(A|B)$ is an increasing function of q, and it is interesting to point out that the decreasing order of the positions of $\sum_{j=1}^{n} S_2(A_j|B_j)$, $\sum_{j=1}^{n} S_1(A_j|B_j)$, $\sum_{j=1}^{n} S(A_j|B_j)$, and $\sum_{j=1}^{n} S_{-1}(A_j|B_j)$ in (2.4) of Corollary 2.4 is quite reasonable since $\sum_{j=1}^{n} S(A_j|B_j) = \sum_{j=1}^{n} S_0(A_j|B_j)$.

§3 Propositions needed to give proofs of the results in §2

By careful scrutinizing nice proofs in [5, Theorem 2.1] and [4, Theorem], we have the following parallel result to [5, Theorem 2.1].

Proposition 3.1. *If f is a continuous, real function on an interval J, the following conditions are equivalent:*

(i) f *is operator concave.*

(ii) $f(C^*AC + t_0(I - C^*C)) \geq C^*f(A)C + f(t_0)(I - C^*C)$

for operator C *with* $||C|| \leq 1$ *and self-adjoint operator* A *with* $\sigma(A) \subseteq J$ *and for fixed real number* $t_0 \in J$.

(iii) $f\Big(\sum_{j=1}^{n} C_j^*A_jC_j + t_0(I - \sum_{j=1}^{n} C_j^*C_j)\Big) \geq \sum_{j=1}^{n} C_j^*f(A_j)C_j + f(t_0)(I - \sum_{j=1}^{n} C_j^*C_j)$

for operators C_j *with* $\sum_{j=1}^{n} C_j^*C_j \leq I$ *and self-adjoint operators* A_j *with* $\sigma(A_j) \subseteq J$ *for* $j = 1, 2, ..., n$ *and for fixed real number* $t_0 \in J$.

(iv) $f(\sum_{j=1}^{n} C_j^*A_jC_j) \geq \sum_{j=1}^{n} C_j^*f(A_j)C_j$

for operators C_j *with* $\sum_{j=1}^{n} C_j^*C_j = I$ *and self-adjoint operators* A_j *with* $\sigma(A_j) \subseteq J$ *for* $j = 1, 2, ..., n$, *where* $n \geq 2$.

(v) $f(PAP + t_0(I - P)) \geq Pf(A)P + f(t_0)(I - P)$

for projection P *and self-adjoint operator* A *with* $\sigma(A) \subseteq J$ *and for fixed real number* $t_0 \in J$.

Corollary 3.2. *If* f *is continuous operator concave function on the half open interval* $[0, \alpha)$ *to* $[0, \alpha)$ *with* $\alpha \leq \infty$, *then*

$$\begin{aligned} f(\sum_{j=1}^{n} C_j^*A_jC_j) &\geq \sum_{j=1}^{n} C_j^*f(A_j)C_j + f(0)(I - \sum_{j=1}^{n} C_j^*C_j) \\ &\geq \sum_{j=1}^{n} C_j^*f(A_j)C_j \end{aligned}$$

for operators C_j *with* $\sum_{j=1}^{n} C_j^*C_j \leq I$ *and self-adjoint operators* A_j *with* $\sigma(A_j) \subseteq [0, \alpha)$ *for* $j = 1, 2, ..., n$.

We recall the following obvious Proposition 3.3.

Proposition 3.3. *Let* $A > 0$ *and* $B > 0$. *Then*

(i) $A\natural_{-1}B = AB^{-1}A$, (ii) $A\natural_2B = BA^{-1}B$, (iii) $A\natural_0B = A$, (iv) $A\natural_1B = B$, *and*

(v) $A\log A \geq \log A$ *for any* $A > 0$.

Remark 3.1. If (i') f is continuous operator concave on J containing 0 and $f(0) \geq 0$, then the following (ii') holds by (i) and (ii) of Proposition 5.1

(ii') $$f(C^*AC) \geq C^*f(A)C + f(0)(I - C^*C) \geq C^*f(A)C$$

for operator C with $||C|| \leq 1$ and self-adjoint operator A with $\sigma(A) \subseteq J$ since $f(0) \geq 0$ and $I - C^*C \geq 0$.

As "*f is continuous operator concave function and $f(0) \geq 0$*" just essentially corresponds to "*f is continuous operator convex function and $f(0) \leq 0$*" in (i) of [5, Theorem 2.1], it turns out that Proposition 3.1 is essentially shown under an additional condition $f(0) \geq 0$ in [5, Theorem 2.1],*briefly speaking, Proposition 3.1 with $f(0) \geq 0$ becomes Theorem 2.1 in* [5].

Remark 3.2. It is shown in [6 ,Theorem 6] that if f is operator monotone function, (iv) of Proposition 3.1 holds. Also Corollary 3.2 implies that if f is an operator monotone function on the half open interval $[0, \alpha)$ to $[0, \alpha)$ with $\alpha \leq \infty$, then $f(\sum_{j=1}^n C_j^* A_j C_j) \geq \sum_{j=1}^n C_j^* f(A_j) C_j$ for operators C_j with $\sum_{j=1}^n C_j^* C_j \leq I$ and self-adjoint operators A_j with $\sigma(A_j) \subseteq [0, \alpha)$ for $j = 1, 2, ..., n$, which is shown in [6 , Corollary 7], bacause f is operator concave on $[0, \alpha)$ to $[0, \alpha)$ with $\alpha \leq \infty$ if and only if f is operator monotone on $[0, \alpha)$ to $[0, \alpha)$ with $\alpha \leq \infty$.

Addendum. After we have written this manuscript, we know that quite similar results to Proposition 5.1 are shown in the following recent paper: F.Hansen and G.K.Pedersem, Jensen's operator inequality, Bull. London Math. Soc., **35**(2003), 553-564.

This paper will appear elsewhere with complete proofs.

References

[1] P.S.Bullen, A dictionary of inequalities, Pitman Monographs and Surveys in Pure and Applied Mthematics 97, LONGRAM, 1998.

[2] J.I.Fujii and E.Kamei, Relative operator entropy in noncommutative information theory, Math. Japonica **34**(1989), 341-348.

[3] J.I.Fujii and E.Kamei, Uhlmann's interpolational method for operator means, Math. Japonica **34**(1989), 541-547.

[4] F.Hansen, An operator inequality, Math. Ann., **246**(1980), 249-250.

[5] F.Hansen and G.K.Pedersen, Jensen's inequality for operators and Löwner theorem, Math. Ann., **258**(1982), 229-241.

[6] M.K.Kwong, Some results on matrix monotone functions, Linear Alg. and Its Appl., **118**(1989), 129-153.

[7] S.Kullback and R.A.Leibler, On information and sufficiency, Ann. Math. Statistics, **22**,1951,79-86.

[8] M.Nakamura and H.Umegaki, A note on the entropy for operator algebra, Proc. Japan Acad., **37**(1961), 149-154.

[9] C.E.Shannon, A mathematical theory of communication, Bull System Technical Journal, **27**(1948),379-423; 623-656.

Takayuki Furuta
Department of Mathematical Information Science, Faculty of Science,
Tokyo University of Science, 1-3 Kagurazaka, Shinjukuku,
Tokyo 162-8601, Japan
e-mail: furuta@rs.kagu.tus.ac.jp